

**Document S4:** *rli* loci in EGD-e.

Sequences of *rli* genes (in bold and red letters) on the corresponding sequences of the intergenic regions (IGRs). Sequences were deduced from 5' RACE data, length observed on Northern blots, folding and rho-independent transcription terminator predictions. Rho-independent terminators are shown for *rliF*, *G*, *H* and *I* (convergent arrows). In few cases (*rliC* and *D*), stem-loops that might be used as transcription terminators are indicated. Sequences provided *rnpB*, *ssrA* and *ssrS* are predicted from the Rfam database (<http://www.sanger.ac.uk/Software/Rfam/index.shtml>). The orientation of the genes is shown by signs "<" and ">". Location of IGRs refers to the annotation of the EGD-e genome (Glaser *et al.* 2001), (see Listilist web site: <http://genolist.pasteur.fr/ListiList/>). Transcription start sites are indicated by "+1" and the base boxed, processing sites are shown by "↓" and the base noted by an upper case underlined. Hexamers -10 (consensus TATAAT) and -35 (consensus TTGACA) of putative RpoD-dependent promoters are boxed.

***rliA* locus:** *lmo0476* < *rliA* > *lmo0477* >

**IGR Location:** 513268-514004 - length : 737 bp

**Function of flanking genes:** *lmo0476* (similar to oxetanocin A resistance protein oxrB);

*lmo0477* (putative secreted protein)

**Comments or specific legend:** *rliA* sequence is deduced from the mapped transcription start site and the length observed on Northern blot using a PCR fragment corresponding to the entire IGR as probe.

```

tcatattcacctctttttttaattatatattagcgaacaaatgttttccatgctttccc
tagcaattccaattttgctgagaaaagttataaaaaattactgaaataaccgaagatgc
gattgataaaatgaaaaccagctcaggtggcggttaggctagttcttggtggcggtgtag
tgtgattattctagacgaaatcccgttggtgtatcatctggctaatacgcccagattgca
catgaaagttgtgggaatttggtttaaggaagtaaacag-35tggaacattccagtcattt
-10tttataaa+1gtaatcccAcggtgataaaaaTacggtttatacacgtgggattatggttcta
tatgaaataggttatttattagaagggtaaatggcttattttatttttcaatgtataaa
attatgtatattgacacaatttggttattttgaacataaatttaagtaaatcaattctt
aatgtattgattttcgcctaaaatgggtcgttcatgacaagaatcggacatttcattaca
tttttggttatacattcattttctatgctatggttttagtaggaataactattttggtac
aggaagtaagcttgtagtacatgtatttccttcattagttgccctgtatgatttacaaa
ttatttttcaaagcatataagccacaacaagtgaataactttaataaatctcatatct
caggaagtgtagctaaa

```



**Locus *rliC* :** *lmo1117* > *rliC* > *lmo01118* <

**IGR Location:** 1154204-1154864 - length: 661 bp

**Function of flanking genes:** *lmo1117* (glyoxalase family protein); *lmo1118* (unknown)

**Comments or specific legend:** *rliC* sequence is deduced from the transcription start site mapped and the length observed on Northern blot using a PCR fragment corresponding to the entire IGR as probe.

```

gatttaaataatatttatgatgaacgataacgctaaaacgacggatgaatgattcgtcg
      -35                -10  rliC> +1 +1
tttttgatg[tttagaa]ataatgcttaagtgggg[taagat]ataaca[ttgtggattaaaata
gatttagtgtat[ttatacactttgttatgaagggttagttaagtattccttgttgaaaaat
tgaaatttaata[ttctagaggcaatatttttagattaaagtgataactttttgataacag
caatttgataagagccaataat[ttgggaaatgaattttaatgcataaaacaacgtc
gaaactaaattaat[gtttaatgtagagttgtccttgctgtgtggaaatagagtgtaat
actcctat[ttttaaggctgaaattgtaatcaagtaaattggttgtccattccttttatt
ttagatatcccttcaagtaaattatcataatttacttgaagggataaa]ttat[ttttatag
      <-----> <----->
ttttaagagagttttat[ttttttacaaatataat[tttaactaaagattcaatg[tttaagtt
gcaaaatag[tcagaatagattt[gtttcaaaattaat[ttatcatc[taatagaaaaaaatt
tcaaatta[aaataatgattaa[ttt[gtggtaaaactat[tttttaca[aaataatcactctgaaa
t

```

**Locus *rliD* :** *rpsO* > *rliD* < *pnpA* >

**IGR Location:** 1359103-1359352 - length: 250 bp

**Function of flanking genes:** *rpsO* (ribosomal protein S15); *pnpA* (polynucleotide phosphorylase, PNPase)

**Comments or specific legend:** *rliD* sequence is deduced from the 5' ends mapped and the length observed on Northern blot using a PCR fragment corresponding to the entire IGR as probe. *rliD* overlaps the 5' end of *pnpA* and is transcribed in the opposite direction. The green box shows the translation start codon of *pnpA*. Note that the most upstream 5' end mapped for RliD is certainly a processing site indicating a transcriptional start further upstream. This locus carries *sraG* gene in *E.coli* and closely relative species (Argaman et al., 2001).

```

aaaaccagacgaaagcgggataccaaaagttcccgcgttttttttgagtattaactatact
aaatTTTggaaaacacgccgaagaaatgaaactttatcgcatgaatccaaaaacgtgctt
caaaaatttgctatTTTTtagcaggTTTTggattcatgtgtgctttagacaatctaaagt
          ↓      ↓
gtgCGTgaacattcttatgcatggtaaagTtctGgttcgtgTTTTcctaatacaaggA
gagtaaaatTgtctgaaaaacaagtTTTTcaacagagtgggcaggtaaaacattatct
          pnpA >
gttgaagtaggtcaattagcaaaacaagcaagtggggcagccttaattcgttacggcgat
acggttgTTTTaacagccgcagtaggttctaaaaaaccacgtccaggcgactTTTTccca
          ↓ <rliD
ttaactGTTaactatgaagaaaaaatgtattccgttggttaaagttcccggtgattctta
aaacgtgaaggacgtccaagcgaccgtgctacattaacagcgcgtcttattgaccgtcca
atccgtccg

```

**Locus *rliE*** : *comC* < *rliE* > *folC* >

**IGR Location:** 1584638-1584850 - length: 213 bp

**Function of flanking genes:** *comC* (similar to *B. subtilis* late competence protein *comC*); *folC* (Folyl-polyglutamate synthetase)

**Comments or specific legend:** *rliE* overlaps the 5' end of *comC* and is transcribed in the opposite direction. No obvious RpoD-dependent promoter can be assigned to *rliE*. The green box shows the translation start site of *comC*.

aaaaagaaaggaaaaaatggggatgatatgataaaaagcaagtttcttttgacaataatt

acactccgagaagcggaataaaaatgatttttctactggaaaacattctcctgcgact **Tg** *rliE*> +1

aatgaaagacataaataccgcactataaatagctaataaaaagtaaat **catctgtcacct**  
<*comC*

cctagttgttaaaaacagaatagcagagacgaaatTTTTTgcacatagctaaaatcact

↓  
TttccatcttctattgcataacaaaaacacctatTTTcattagtggaaaacaggtgtttt  
← tatatTTTTatcttctgaaaatcTTTTtagctatTTgttctTTTTtatttctcacag

tatactcgatttattcaaaagta

**Locus *rliF* :** *nadA* > *rliF* < *lmo2026* >

**IGR Location:** 2106001-2106329- length: 329 bp

**Function of flanking genes:** *nadA* (quinolinate synthetase); *lmo2026* (putative peptidoglycan bound protein with LPXTG motif)

**Comments or specific legend:** No 5' end was mapped for *rliF*. The sequence provided is speculated from folding analysis and length observed on Northern blots.

```
ataaaaaaaaaacgaaagcaacaactgctttcgtttcagctctataatttataaaagcttct
at ttgtcaatct aaat t t t a a t c t g t a g a a c t t a c t c g a g t a t a t g c t a c t a t c a t g a c t
aaa a a g c a a c t c t t t t c a c a a c c a a t t a a t a t c c a c c c t a g c t a t a c a t t g a a a a a g t g
t a t t t a c a a t g a t g g t t t t g g g t a t a t t t t g g g t a t a c c c t t t t t t t a c a t t t t t g c t t g
cc t t g a t t t a a a t a a c a a a t t t c t c t t a a a t a a a a c t a g g t t g c c c a t c <rliF t t t t t t c t
caagtgggcaacctagttttatttatattt
```

**Locus rliG:** *lmo2302* < *rliG* < < *lmo2303*

**IGR location:** 2386680-2387046 - length: 367 bp

**Function of flanking genes:** *lmo2302* (unknown); *lmo2303* (Gp66 protein of the A118 phage)

```

tttgaacaccctttttatcttctactataacaaataaaaccacctgctcaatcttcaac
←
agatggaaagggtatatatcttaaaaaactggttaacgcaccagtcagcgccacatgcgt
gttttacatccagtatggataggatatgagattgaacagaagtgtcgtcatctgttgaga
ctagtgccagatacaaagcctctgccgggcaaCatagcaatctcctgctatatcatcat
aagattataaatgagaagtggagcgcagactcaatataagattttatctttgtaatcatc
+1 <rliG -10 -35
ttcacttctcacTaaataacattttatcaccttttttgctcaaaaagtgccagaaaagtg
ccatctt

```



**Locus *rliI*:** *lmo2760* < *rliI* > *lmo2761*

**IGR location:** 2841896 – 2842413 - length: 518 bp

**Function of the flanking genes:** *lmo2760* (ABC transporter); *lmo2761* (beta-glucosidase)

```

ttttatcctccgtttatccaaggaagcatggcaaaatgtttcacgtgaaacatTTTTTca
gacaacAAAAAAGCGTcagacaagCGCCGACGCCAATTATAAGTATAATACTAATTTAG
      ↑                ↓
cactatacaaaccgattgcgtgactccatgcttctctcattttttctcgttatacagttt
      ↓                ↓                ↓
tttgggcagacttcgcaaTgtaactgcaatAccggaaAaaattgcacggacaatataca
      ↓                ↓                ↓
aaaatttagctatactagttcgcgcatcAAAAGAAACATGTCATCTCACCTCTTTTCAT
      +1 <rliI -10                -35
taatAcatatagaataaaggatgctgccccaaaTgtcaattggttctccaatcacatgca
cgaaacacagcgtttcaactcagaaacttataccattgatgtgaaacacgaacccaaagc
acaaaaatctattttaccaattagtcgaaaccgtttacattaataaggcaagcaaaaactaa
aacagactaatcgcacataatTTTTTGGAGGTGCCATAA
    
```

**Locus *rnpB*:** *lmo1887* < *rnpB* < < *lmo1888*

**IGR location:** 1961644 –1962225 - length: 589 bp

**Function of the flanking genes:** *lmo1887* (putative RNA methylase); *lmo1888* (unknown)

```
ttgtgtcccatcctttaattaatttctaataaaaaagctgttttaagag
ggggagggtttgttttaaataggcattcaaccagggtcgaacaccaccttaaaacttgcg
tccggccttctaaaacagccttcaatgttaaattaaatcctgtaaataacgctctgtaag
ccatgttctgttctttactacgtacagggtaatagtaaagagataaccatctatctgcg
aaaaatacatttgcgctcgtcccttggttcttgattcctcgggaaaagtgccctacat
aagtttgagtttctcgctcgtggggttacctcgttccacttccgagcatttctgccgga
acttcgtcactgtggcaccttcaaggttataagaccatatccgaagacttaggtcattca
cctgccgtcaaaccgaaaaccggaatgcctggcttattgtttgaccaggcacgatcac
tacgggcatcacagcaccgtgcgagcatggactttcctctgctttcaaaagaaaacagcg
attatccgaacgttattcatttatttctaacttacttatactagcatacctattcgt
<rnpB
```

**Locus *ssrS*:** *lmo2443* < *ssrA* < < *lmo2444*

**IGR location:** 2509411 –2510361 - length: 950 bp

**Function of the flanking genes:** *lmo2443* (unknown); *lmo2444* (glycosyl hydrolase)

aagcaaacctcctcctttggtacatattagcataaaaagttaatgtataaaaaactgct  
caagaatggcttttcttcgttttaataagcaaaaaatggtttttatagtaattatta  
cactttatgtggtgaatttgctcttttagataaatttcacacctatatgtggcgttat  
tcaaaaattattcagtttgtaaatacaaacagcaaaaagcgccataaacattcatttaa  
actattcatttatcattttttattaatttttagtagtaagaaaaaattacatttgaatg  
gatatgtatatttattcagtcacatggttatttctaaaatcgtaattgttttcttgactc  
catcgttttagcacccaataaaaaagtgatgcaaacagcttaacaagcagcttgatca  
cttttcattatccataaaaata**tgagatggtgggagtcgaaccacgtccagaaatc**  
**ggcacttaaatatctacgagcgtagtcaccgtattagcatttcgcataaacatcggcca**  
**gtaacaggcttcattcagctagtctgattaagctcttctatttaacccagacggagat**  
**taactagtgtagcccacttagagtgagacccttaccgtagcacatgggcatgcacggag**  
**gatcagctatgctactgcttatttaggcagcgaagctaggttttgttttctttgccagt**  
**tattattggctttgacggttgataacgaggacgatcccccgactcgcaactcaagctcga**  
**actatccctgtcgaatccgtaacatcccc**agagttaaagtctaacagtagttattataca  
**<ssrA**  
ggaaaaaacgaggctttacaacagaaaaacttgtgacatagccaaatgcgcaactatt  
athtagcatgtcacaagtttccattaattaatggttattttactgaat

**Locus *ssrS*:** *lmo1513* < *ssrS* < < *lmo1514*

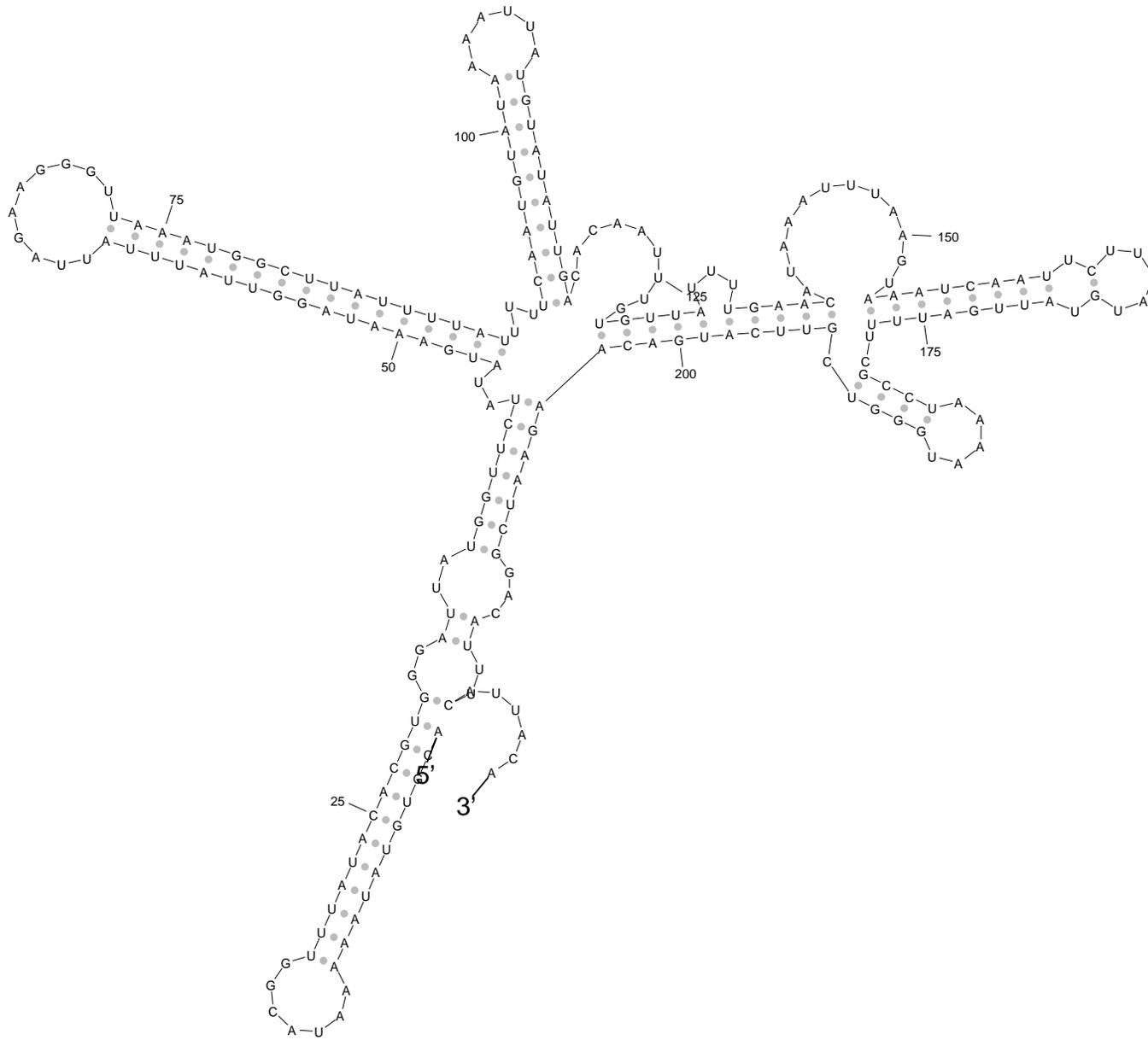
**IGR location:** 1546222–1546608 - length: 387 bp

**Function of the flanking genes:** *lmo1513* (iron-sulfur cofactor synthesis); *lmo1514* (ATPase, AAA family domain)

```
ttttttatcgcgtccttttttagacgaattcctggttcgccgtaaatcgttt
cttattgtaacagcaataactcctgtaaagtaaagaggaataactcctatttacacgcaa
aaagaaacccaatcgtaccgtttaggattcccttcctttgaaccgctctcaattaaag
caggtgggtgctctgtctaacggtttgacttcctcgtgtgaaggcatgtgcgccgccc
ggaactccaagctccctaggtataaaaaattgttcggttcaaaaatcaggaactatat
cgaatacattagggtttcttttctttatttagaataccataggaagacgaaatttcaac
<ssrS
gtttagcgacttttcatttggcactttttttattct
```

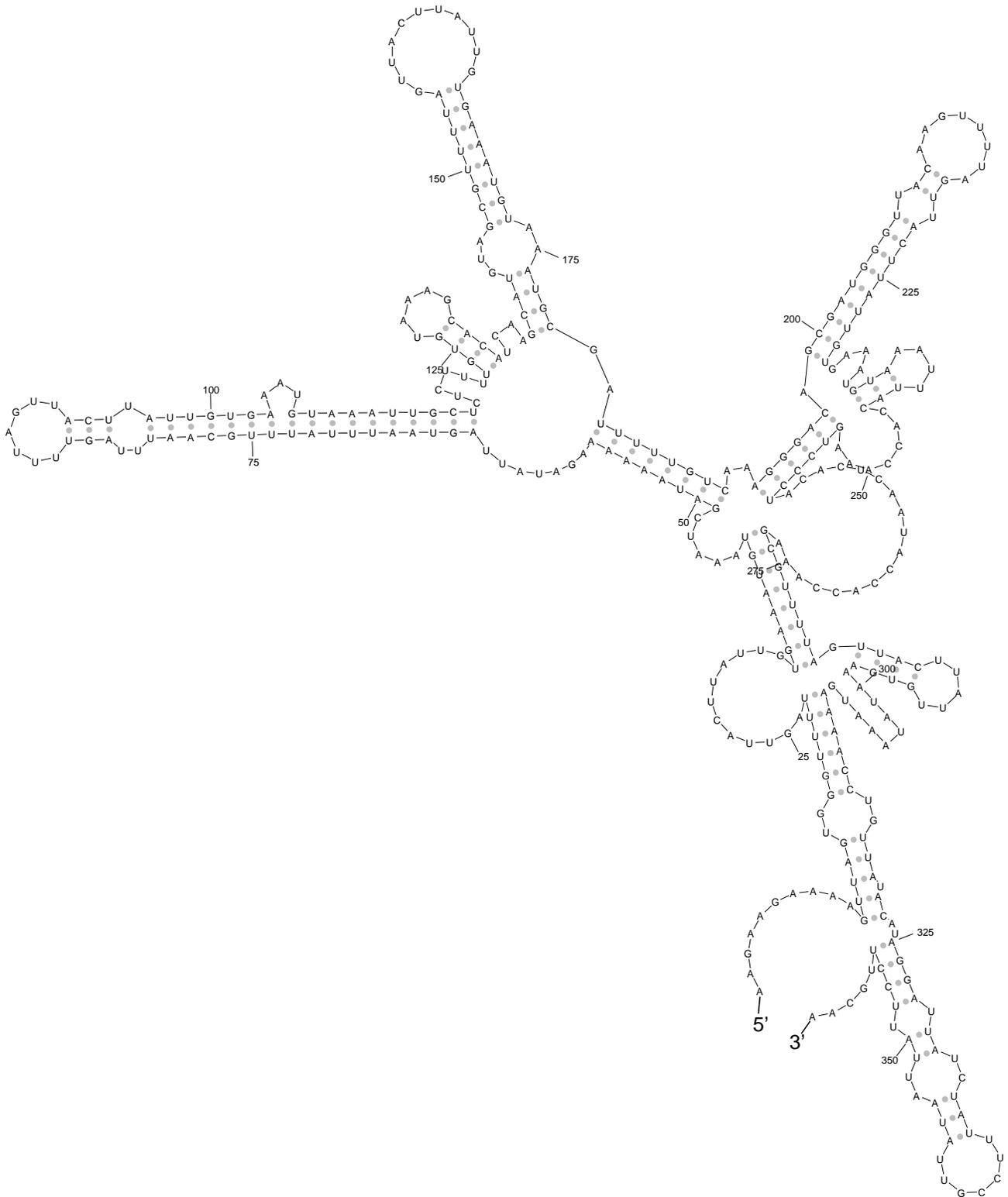
**Document S5** : folding predictions of Rli ncRNAs by MFOLD program.

# RliA



$$\Delta G_{37}^{\circ} = -53.90$$

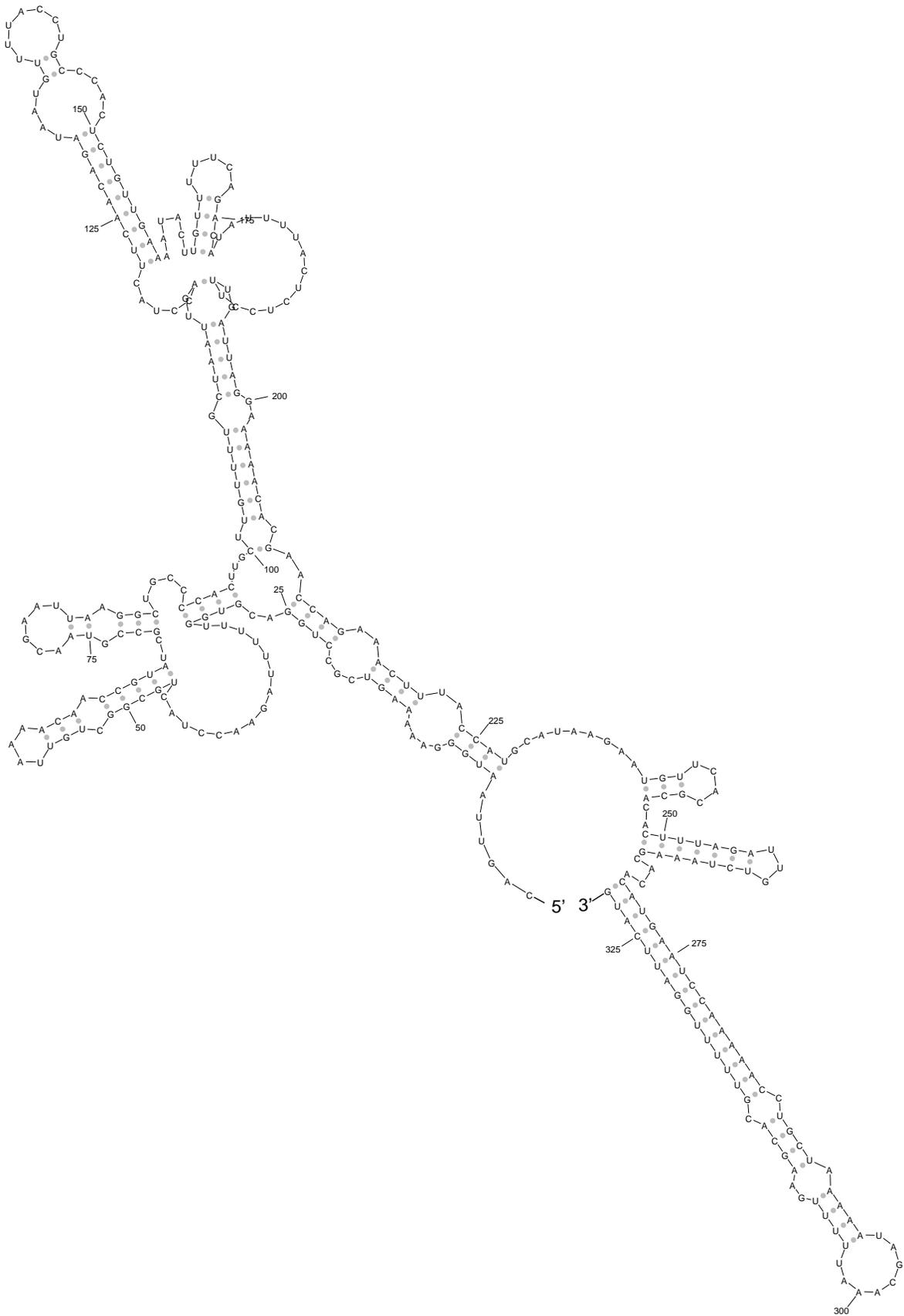
# Riib



$$\Delta G_{37}^{\circ} = -69.60$$

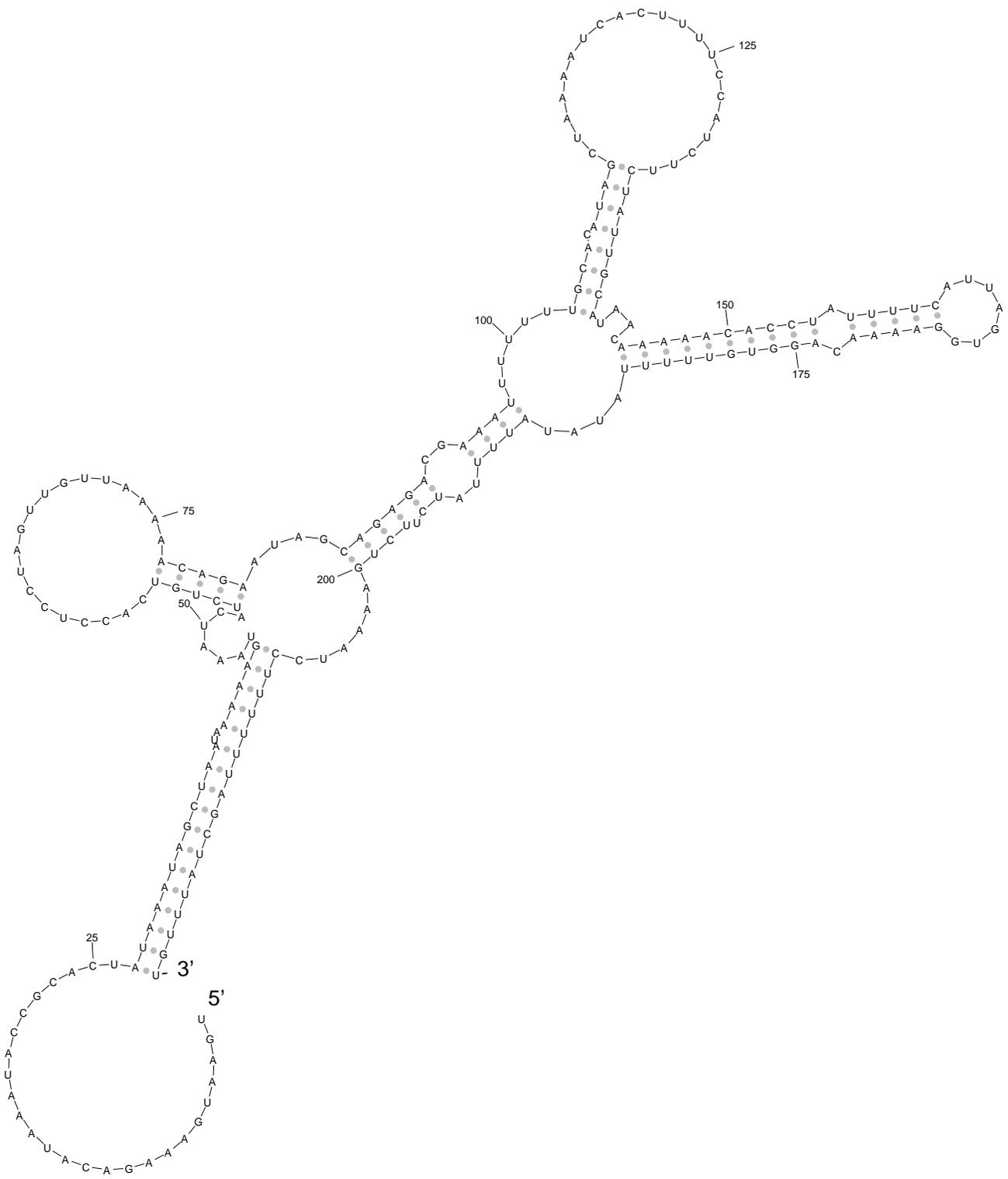


# RiD



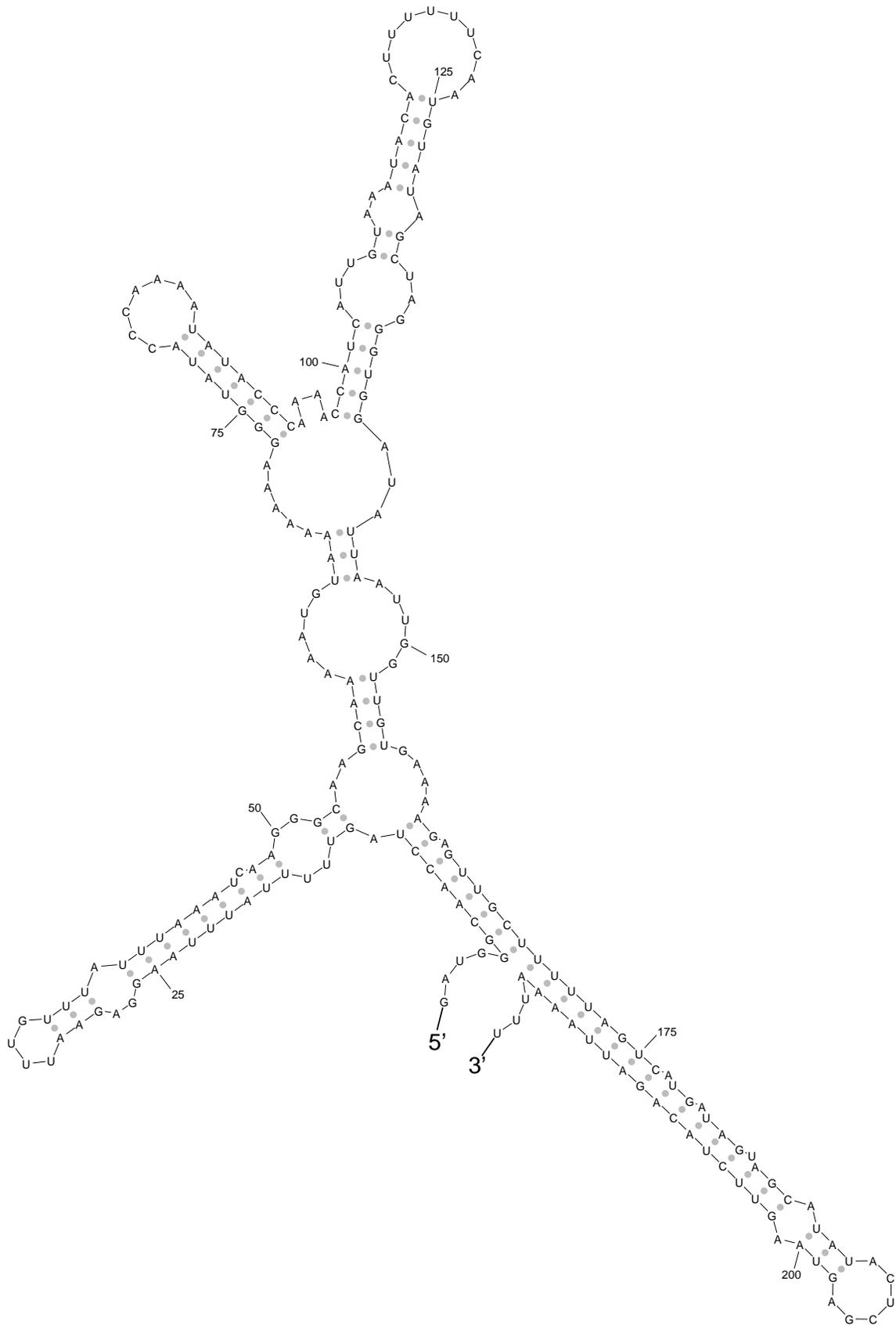
$$\Delta G_{37}^{\circ} = -77.20$$

# RliE



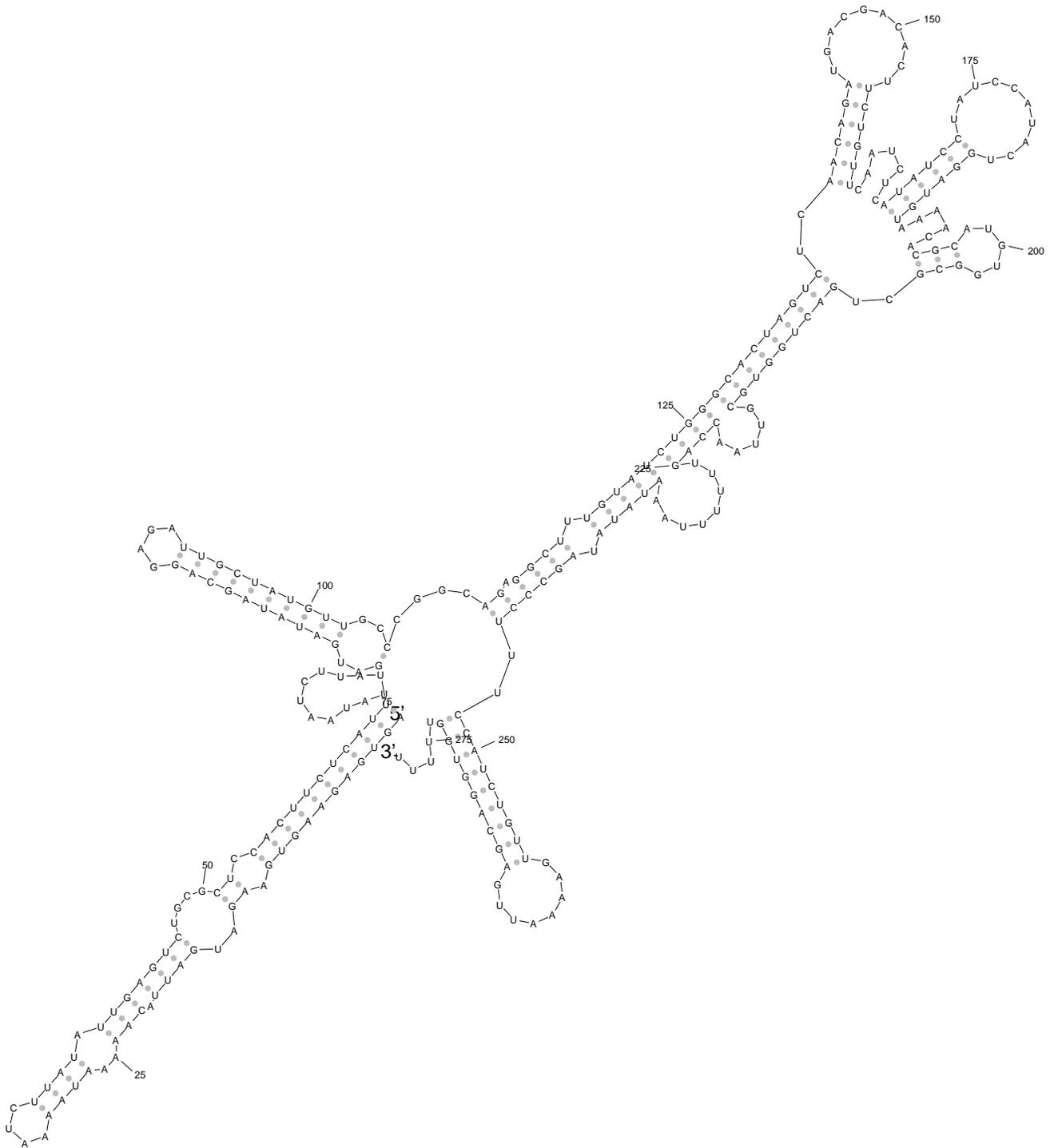
$$\Delta G_{37}^{\circ} = -40.40$$

# Rif



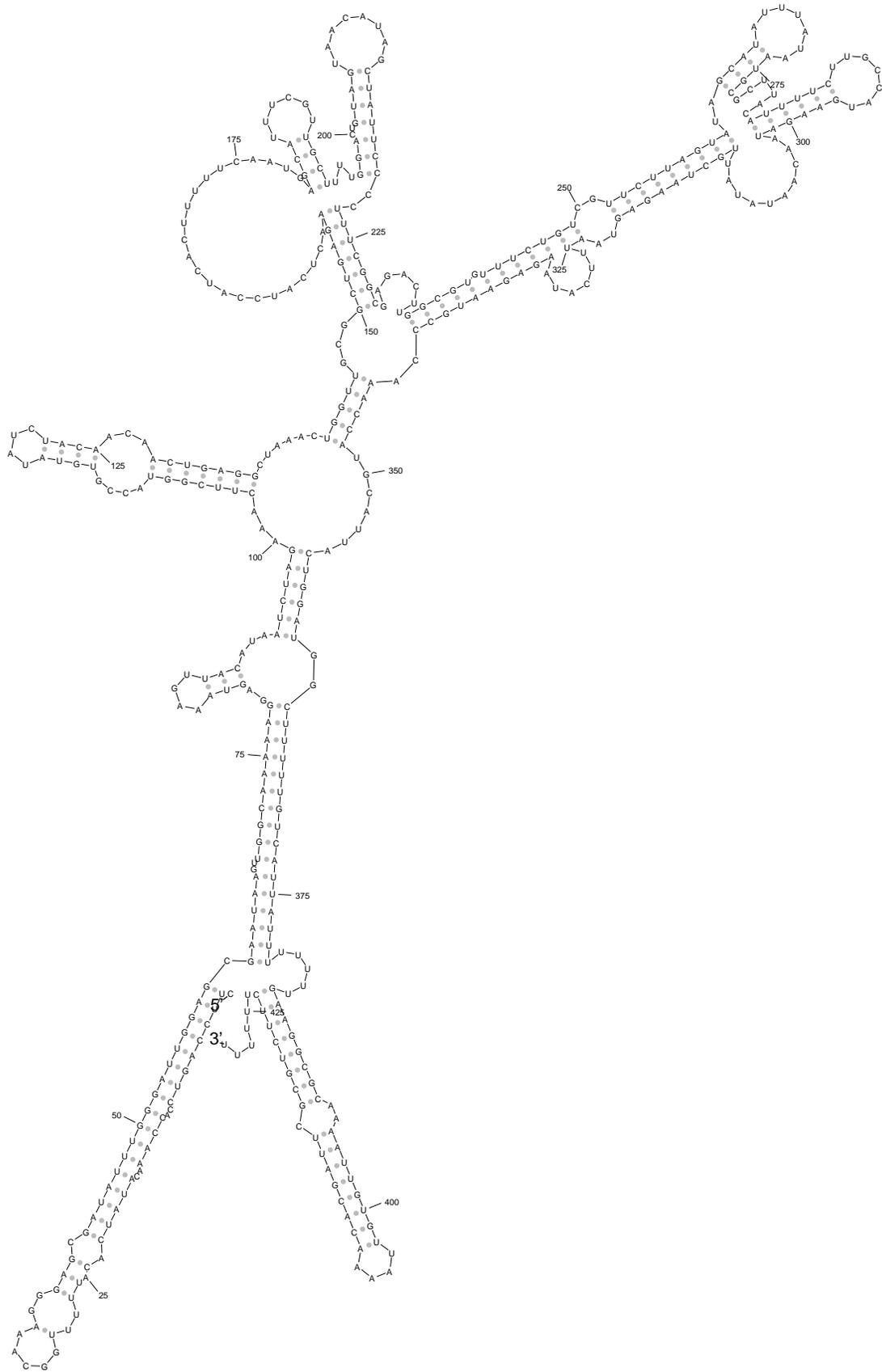
$$\Delta G_{37}^{\circ} = -41.40$$

# RliG



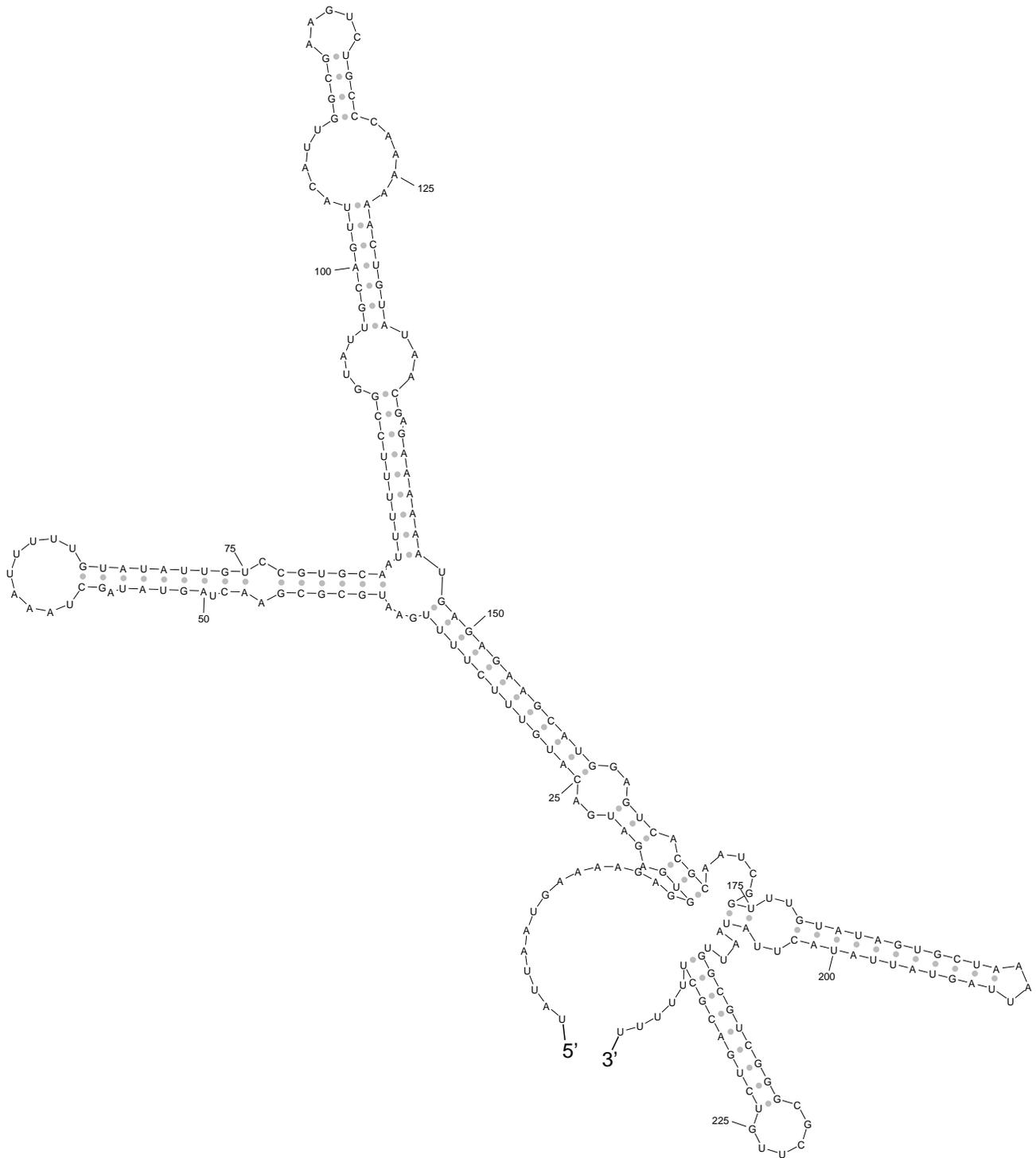
$$\Delta G_{37}^{\circ} = -77.00$$

# R11H



$$\Delta G_{37}^{\circ} = -107.60$$

# Ril



$$\Delta G_{37}^{\circ} = -77.80$$